

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 9, 2002, 00:09:19 ; Search time 8498.8 Seconds

(without alignments)
31.610 Million cell updates/sec

Title: US-09-851-670-12

Sequence: 1 acagctgcgcccaataacataatc 25

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 111874

Minimum DB seq length: 0

Maximum DB seq length: 60

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

EST: *
1: em_estfun:*
2: em_estfun:*
3: em_estfun:*
4: em_estfun:*
5: em_estfun:*
6: em_estfun:*
7: em_estfun:*
8: em_estfun:*
9: em_estfun:*
10: em_estfun:*
11: em_estfun:*
12: em_estfun:*
13: em_estfun:*
14: em_estfun:*
15: em_estfun:*
16: em_estfun:*
17: em_estfun:*
18: em_estfun:*
19: em_estfun:*
20: em_estfun:*
21: em_estfun:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	DB ID	Description
1	13.4	53.6	33 13	TA80C11Q
2	13.4	53.6	50 10	TA80C11Q
3	13.4	53.6	50 10	TA80C11Q
4	13.4	53.6	50 10	TA80C11Q
5	13.4	53.6	50 10	TA80C11Q
6	13.4	53.6	50 10	TA80C11Q
7	13.4	53.6	50 10	TA80C11Q
8	13.4	53.6	50 10	TA80C11Q
9	13.4	53.6	50 10	TA80C11Q
10	13.4	53.6	50 10	TA80C11Q
11	13.4	53.6	50 10	TA80C11Q
12	13.4	53.6	50 10	TA80C11Q

13	12.6	50.4	50	10	AU104819	AU104819
14	12.6	50.4	60	13	A2460727	A2460727
15	12.4	49.6	36	11	A2331596	A2331596
16	12.4	49.6	55	11	BG152186	BG152186
17	12.4	49.6	57	13	A2648926	A2648926
18	12.2	48.8	28	13	TA380E12Q	TA380E12Q
19	12.2	48.8	39	13	A2986277	A2986277
20	12.2	48.8	40	13	AA588339	AA588339
21	12.2	48.8	40	13	A2665717	A2665717
22	12.2	48.8	50	10	AU102603	AU102603
23	12.2	48.8	50	10	AU104015	AU104015
24	12.2	48.8	50	10	AU104017	AU104017
25	12.2	48.8	51	13	TA289F12Q	TA289F12Q
26	12.2	48.8	53	13	B36066	B36066
27	12.2	48.8	58	10	A1499235	A1499235
28	12.2	48.8	59	10	A1954847	A1954847
29	12.2	48.8	28	13	A2485440	A2485440
30	12.2	48.8	31	11	C01612	C01612
31	12.2	48.8	41	13	A2484492	A2484492
32	12.2	48.8	45	13	A2814934	A2814934
33	12.2	48.8	50	13	A2659610	A2659610
34	12.2	48.8	50	13	A2918083	A2918083
35	12.2	48.8	53	13	A2807379	A2807379
36	12.2	48.8	54	10	AA855787	AA855787
37	12.2	48.8	54	13	B03584	B03584
38	12.2	48.8	54	13	TA374D02P	TA374D02P
39	12.2	48.8	57	13	A2450809	A2450809
40	12.2	48.8	58	11	N28963	N28963
41	12.2	48.8	59	10	A1337657	A1337657
42	11.8	47.2	23	13	A2339893	A2339893
43	11.8	47.2	39	13	A2331190	A2331190
44	11.8	47.2	43	10	A1461342	A1461342
45	11.8	47.2	49	13	A2810303	A2810303

ALIGNMENTS

RESULT 1	TA80C11Q	33 bp	DNA	GSS	13-DEC-2000
LOCUS	TA80C11Q				
DEFINITION	T. brucei sheared genomic DNA clone 80c11, reverse sequence.				
ACCESSION	AL461346				
VERSION	AL461346.1				
KEYWORDS	GENE				
SOURCE	Trypanosoma brucei				
ORGANISM	Trypanosoma brucei				
REFERENCE	Trypanosoma				
AUTHORS	Hall, N., Bowman, S., Lennard, N.J., Doggett, J., Atkin, R., Chillingworth, C., Ormond, D., Harris, B., El-Sayed, N., Hou, L., Melville, S.E., Rajandream, M.A. and Barrell, B.G.				
TITLE	Direct Submission				
JOURNAL	Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton, Cambridgeshire CB10 1SA, E-mail: barrell@sanger.ac.uk and nh@sanger.ac.uk				
COMMENT	Constructed at the Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from a cloned population of Trypanosoma brucei (FRE927/4 GUTat 10.1) was mechanically sheared to give a tight size distribution (4 kb). The v + i method used for the library construction is described in detail in Smith, H. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999).				
FEATURES	Email: nelsayed@tigr.org				
SOURCE	Details of T. brucei sequencing at the Sanger Centre are available at http://www.sanger.ac.uk/Projects/T_brucei/				
	Location/Qualifiers				
	1..33				

```

Query Match
Best Local Similarity 53.6%; Score 13.4; DB 13; Length 33;
Matches 17; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 cagctcgccccatcaatct 24
      ||||| || | | || |
Db 38 CAGCTCGCTCCACTGACACACT 4

RESULT 2
AUI03898 50 bp mRNA EST 05-APR-2001
LOCUS AUI03898/c
DEFINITION KAT01688 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
AUI03898
ACCESSION AUI03898.1 GI:13553419
VERSION AUI03898.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 50)
Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hatae,
H., Ota,T., Iisaga,T., Tanaka,T., Nakamura,Y., Morishita,S., Okudo,
K., Suyama,A. and Sugano,S.
Fine structural analysis of transcription start sites of human
mRNAs using full-length enriched and 5'-end enriched cDNA libraries
unpublished (2001)
Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ytsuzuki@ims.u-tokyo.ac.jp
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano,
S. Construction and characterization of a full length-enriched and
a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
Location/Qualifiers
1..50
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="KAT01688"
/clone_lib="Sugano Homo sapiens cDNA library"

BASE COUNT 11 a 11 c 20 g 8 t
ORIGIN

FEATURES
Source
1..50
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="KAT01688"
/clone_lib="Sugano Homo sapiens cDNA library"

Query Match
Best Local Similarity 53.6%; Score 13.4; DB 10; Length 50;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 cagctcgccccatt 16
      ||||| ||||| |
Db 38 CAGCTCGCCCCCAGT 24

RESULT 3
AUI07252 50 bp mRNA EST 05-APR-2001
LOCUS AUI07252/c
DEFINITION LNC08445 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
AUI07252
VERSION AUI07252.1 GI:13556773
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

```

	REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
	AUTHORS	I (bases 1 to 50)
TITLE	Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hatake,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo,K., Suyama,A. and Sugano,S.	
JOURNAL	File Structural analysis of transcription start sites of human mRNAs using full-length enriched and 5'-end enriched cDNA libraries	
COMMENT	Unpublished (2001) Contact: Yutaka Suzuki Department of Virology Institute of Medical Science, University of Tokyo 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan Email: yusuzuki@ims.u-tokyo.ac.jp Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Suganoo,S. Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).	
FEATURES	location/qualifiers	
SOURCE	1..50 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="LINC08445" /clone_lib="Sugano Homo sapiens cDNA Library"	
BASE COUNT	18 a 9 c 16 g 7 t	
ORIGIN		
Query Match	52.0%; Score 13; DB 10; Length 50;	
Best Local Similarity	76.2%; Pred. No. 7.3e+04;	
Matches	16; Conservative 0; Mismatches 5; Indels 0; Gaps 0;	
OY	5 ctgcgcccaattacatatc 25 Db 43 CTGGGCCCATTAACACCTCC 23	
RESULT	4	
LOCUS	w71856	
DEFINITION	w71856 52 bp mRNA EST 18-JUN-1996 me45f07.r1 Soares mouse embryo NbMEI3.5 14.5 Mus musculus CDNA clone IMAGE:390469 5' similar to SW:NDPM_BOVIN P42029	
ACCESSION	NMDH-BIQUINONE OXIDOREDUCTASE 19 KD SUBUNIT ;, mRNA sequence. w71856	
VERSION	w71856.1 GI:1381943	
KEYWORDS	EST.	
SOURCE	house mouse. Mus musculus	
ORGANISM	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. I (bases 1 to 52) Mairra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubouque,T., Geiselt,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Treising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R. The WashU-HMI Mouse EST Project Unpublished (1996) Contact: Maria M/Mouse Est Project WashU-HMI Mouse EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel.: 314 286 1800 Fax: 314 286 1810 Email: mouseest@watson.wustl.edu This clone is available royalty-free through LMNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. NCBI:242301 Trace considered overall poor quality Possible reversed clone: similarity on wrong strand Seq primer: mob.REGA+ET High quality sequence stop: 1. Location/Qualifiers 1..52	
FEATURES		
SOURCE	1..52	

```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:390469"
/clone_lib="Soares mouse embryo NbME13.5 14.5"
/sex="unknown"
/tissue_type="embryo"
/dev_stage="13.5-14.5dpc total fetus"
/lab_host="DH10B"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTACCAATCTGAGAGTGGAGCGCCGGAATTTTGTGTGTGTGTGT
T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
14.5dpc embryos [total RNA provided by Minoru Ko, Wayne
State Univ., from 2 1]; double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pT73 vector. Library went through one round of
normalization, and was constructed by Bento Soares and
M. Fatima Bonaldo."
BASE COUNT      14 a      15 c      15 g      8 t
ORIGIN

Query Match      52.0%; Score 13; DB 11; Length 52;
Best Local Similarity 76.2%; Pred. No. 7.3e+04;
Matches 16; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      1 acagctgcgcccatcaacat 21
        ||||| 1 ||||| 1 |||
Db      23 ACAGCTGAGCGCCATCATCAT 43

RESULT 5
LOCUS   BF643572      52 bp      mRNA      EST      20-DEC-2000
DEFINITION   NF057E06E3F1051 Elicited cell culture Medicago truncatula cDNA
ACCESSION   BF643572
VERSION     BF643572.1 GI:11908793
KEYWORDS    EST.
SOURCE      barrel medic.
ORGANISM    Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eustosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
1 (bases 1 to 52)
Torres-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,
Flores,H.R., Inman,J.T., Weller,J.W. and May,G.D.
Expressed Sequence Tags from the Samuel Roberts Noble Foundation -
Center for Medicago Genomics Research
Unpublished (2000)
Contact: Dixon RA
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7302
Fax: 580 221 7380
Email: radixon@noble.org
Insert Length: 52 Std Error: 0.00
Plate: 057 row: E column: 06
Seq primer: TCACACAGAGAAACAGCTATGAC.
Location/Qualifiers
1..52
/organism="Medicago truncatula"
/db_xref="taxon:3880"
/clone="NF057E06EC"
/clone_lib="Elicited cell culture"
/tissue_type="Cell cultures derived from root tissues"
/dev_stage="Cell suspensions were subcultured every 14
days. Cells were induced six days after subculture"
FEATURES
source

```

```

/notes="Vector: Lambda Zap; Cells were induced with yeast
cell wall extracts equivalent to 50ug/ml glucose in the
final concentration. Samples were taken at 0.5, 1, 12 and
24 hours after induction. Equal amounts of RNA from each
time point were pooled and used for mRNA isolation."
BASE COUNT      9 a      16 c      3 g      24 t
ORIGIN

Query Match      52.0%; Score 13; DB 11; Length 52;
Best Local Similarity 76.2%; Pred. No. 7.3e+04;
Matches 16; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      5 ctgcgcccatcaacatc 25
        ||| 1 ||| 1 ||||| 1 |
Db      31 CTCCTCTTACACATATATC 51

RESULT 6
LOCUS   TA235H01P      52 bp      DNA      GSS      13-DEC-2000
DEFINITION   T. brucei sheared genomic DNA clone 235h01, forward sequence,
genomic survey sequence.
ACCESSION   AL481430
VERSION     AL481430.1 GI:11847124
KEYWORDS    GSS.
SOURCE      Trypanosoma brucei.
ORGANISM    Trypanosoma brucei
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
Trypanosoma.
1 (bases 1 to 52)
Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
McVillie,S.E., Rajandream,M.A. and Barrell,B.G.
Direct Submission
Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
Cambridge CB10 1SA. E-mail: barrell@sanger.ac.uk and
nh@sanger.ac.uk
Constructed at the Institute for Genomic Research (TIGR),
Rockville, MD. Genomic DNA isolated from a cloned population of
Trypanosoma brucei (FRE927/4 GUTat 10.1) was mechanically sheared
to give a tight size distribution (
4 kb). The v + i method used for the library construction is
described in detail in Smth, H. and Venter, J.C. (Making small
insert libraries for whole genome shotgun sequencing projects. In
Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
Barrell, Oxford University Press, 1999).
Email: nelsayed@tigr.org
Details of T. brucei sequencing at the Sanger Centre are available
at http://www.sanger.ac.uk/Projects/T_brucei/.
Location/Qualifiers
1..52
/organism="Trypanosoma brucei"
/strain="FRE927"
/db_xref="taxon:5691"
/clone="235h01"
BASE COUNT      17 a      16 c      9 g      10 t
ORIGIN

Query Match      52.0%; Score 13; DB 13; Length 52;
Best Local Similarity 76.2%; Pred. No. 7.3e+04;
Matches 16; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      5 ctgcgcccatcaacatc 25
        | |||| 1 ||||| 1 |||
Db      19 CACGCCATCTTATACATATATC 39

RESULT 7
LOCUS   AT723111      49 bp      mRNA      EST      07-JUN-2001

```

DEFINITION fc25g02.y1 zebrafish wasnu mpimg EST Danio rerio cDNA clone
IMAGE:3722450 5' similar to SW:HIRA_FUGRU 042611 HIRA PROTEIN ;
RNA sequence.
ACCESSION A1723111
VERSION A1723111.1 GI:5041440
KEYWORDS EST.
SOURCE zebrafish.
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Rasbora; Danio.
REFERENCE 1 (bases 1 to 49)
AUTHORS Clark, M., Johnson, S.L., Lehnach, H., Lee, R., Li, F., Marra, M., Eddy
S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood
K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B.,
Walker, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Rittler, E.,
Kohn, S., Shih, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R.
and Wilson, R.
Wasnu zebrafish EST Project 1998
Unpublished (1998)
Other ESTs: fc25g02.x1
Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbrafish@wustl.edu
CDNA Library Preparation: Matthew Clark, CDNA Library Arrayed by:
Matthew Clark. DNA Sequencing by: Washington University Genome
Sequencing Center Clone distribution: Genome Systems, St. Louis,
Missouri (web address: www.genomesystems.com) (email contact:
info@genomesystems.com) and Research Genetics, Huntsville, Alabama
(web address: www.resgen.com) (email contact: info@resgen.com) and
Resourcenet/centrum/primatdatenbank, Berlin, Germany (web address:
www.rzpd.de)
Trace considered overall poor quality
Seq primer: T3 RT from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1..49
/organism="Danio rerio"
/db_xref="taxon:7955"
/clone_image="3722450"
/clone_lib="zebrafish wasnu mpimg EST"
/sex="mixed"
/tissue_type="26 somite embryos, adult livers, shield
stage embryos"
/lab_host="X11-blue MRF"
/note="Vector: pSPORT1; Site 1: NotI; Site 2: SalI; 1st
strand cDNA was primed with a Not I - oligo(dT)15 primer
(5'/GACTAGTCTAGATCGGACGCGCCCTTTTCTTTTCTTTT3-);
double-stranded cDNA was ligated to Sal I adaptors (BRL),
digested with Not I and cloned into the Not I and Sal I
sites of the pSPORT1 vector (BRL). Library was constructed
by Matthew Clark (Lehnach lab; ICGF, London and Max Planck
Institut fuer Molekulare Genetik Berlin). cDNAs for EST
hybridization were selected following oligonucleotide
zebrafish late somitogenesis (26 ss), adult liver or
embryonic shield stage (5.6 h) libraries. Fingerprint
data were used to computationally cluster cDNAs, and a
single cDNA from each cluster was chosen for sequencing.
In some cases multiple members of the same cluster were
sequenced to assess clustering parameters or single clones
were sequenced additional times to assess quality
control."

BASE COUNT 15 a 14 c 5 g 15 t
ORIGIN

Query Match 51.2%; Score 12.8; DB 10; Length 49;
Best Local Similarity 70.8%; Pred. No. 8.9e+04;

Matches 17; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Qy 1 acagctgcgccattacatt 24
||||| 1111 1111
Db 1 ACAGCTGCCACACAGATTCTTTT 24

RESULT 8
LOCUS AUI07456 50 bp mRNA EST 05-APR-2001
DEFINITION AUI07456 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
HS100655, mRNA sequence.
ACCESSION AUI07456
VERSION AUI07456.1 GI:13556977
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 50)
AUTHORS Suzuki, Y., Tsunoda, T., Taira, H., Mizushima-Sugano, J., Seese, J., Hata
H., Ota, T., Isogai, T., Tanaka, T., Nakamura, Y., Morishita, S., Okubo
K., Suyama, A. and Sugano, S.
Fine Structural analysis of transcription start sites of human
mRNAs using full-length enriched and 5'-end enriched cDNA libraries
Unpublished (2001)
Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp
Suzuki, Y., Yoshitomo-Nakagawa, K., Maruyama, K., Suyama, A. and Sugano
S. Construction and characterization of a full length-enriched and
a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
Location/Qualifiers
1..50
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="HS100655"
/clone_lib="Sugano Homo sapiens cDNA library"

BASE COUNT 12 a 14 c 14 g 10 t
ORIGIN

Query Match 51.2%; Score 12.8; DB 10; Length 50;
Best Local Similarity 70.8%; Pred. No. 8.9e+04;
Matches 17; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Qy 1 acagctgcgccattacatt 24
||||| 1111 1111
Db 1 ACAGCAGACACAGATTGACTTAT 37

RESULT 9
LOCUS AAI90304 52 bp mRNA EST 17-FEB-1997
DEFINITION AAI90304 mi94h01.x1 Soares mouse 3MDMS Mus musculus cDNA clone IMAGE:637585
5' similar to SW:CATD_CHICK Q05744 CATHEPSIN D PRECURSOR ; mRNA
sequence.
ACCESSION AAI90304
VERSION AAI90304.1 GI:1778984
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 52)
AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
The wasnu-HMIT Mouse EST Project

JOURNAL
COMMENT

Unpublished (1996)
Contact: Maria M/Mouse Est Project
Washu-HMI Mouse Est Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:389577

FEATURES

SOURCE

1. .52
/organism="Mus musculus"
/strain="C57BL/6J"
/db.xref="taxon:10090"
/clone="IMAGE:637585"
/clone_id="Soares mouse 3nbms"
/sex="male"
/tissue_type="Spleen"
/dev_stage="4 weeks"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTACCAATCTGACGAGCGCGCGCGCTGTGTGTGTGTGTGTGT
3']; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. RNA
provided by Dr. Bertrand Jordan. Library went through
three rounds of normalization, and was constructed by
Bento Soares and M.Fatima Bonaldo."

BASE COUNT
ORIGIN

10 a 10 c 22 g 10 t

Query Match

51.2%; Score 12.8; DB 10; Length 52;

Best Local Similarity 87.5%; Pred. No. 9e+04; 2; Indels 0; Gaps 0;

Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 cagctgcgcccatat 17

DB 31 CAGCTCTCCCATCA 16

RESULT 10

AZ662142/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AZ662142 30 bp DNA GSS 14-DEC-2000
IM0541M03F Mouse 10kb plasmid UGCG1M library Mus musculus genomic
clone UGCG1M0541M03 F, DNA sequence.
AZ662142
AZ662142.1 GI:11799288
GSS.
house mouse.
Mus musculus
Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 30)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamll,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT

FEATURES

SOURCE

84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0541 row: M column: 03
Seq primer: CGTGTAAACGACGCGCAGT
Class: plasmid ends
High quality sequence stop: 30.
Location/Qualifiers
1. .30
/organism="Mus musculus"
/strain="C57BL/6J"
/db.xref="taxon:10090"
/clone="UGCG1M0541M03"
/clone_id="Mouse 10kb plasmid UGCG1M library"
/sex="male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PMD42 (914732141gb/AP129072.1), a copy-number
inducible derivative of plasmid RI. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
ORIGIN

10 a 4 c 6 g 10 t

Query Match

50.4%; Score 12.6; DB 13; Length 30;

Best Local Similarity 78.9%; Pred. No. 1e+05; 4; Indels 0; Gaps 0;

Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 6 tcgcccattacatt 24

DB 27 TGGCCACATTAATAAT 9

RESULT 11

AA857265/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AA857265 49 bp mRNA EST 09-MAR-1998
of75f01.st NCI CGAP Co8 Homo sapiens cDNA clone IMAGE:1436185 3'
similar to SW:UIC6_HCVWA P16836 HYPOTHETICAL PROTEIN U1.26. ;, mRNA
sequence.
AA857265
AA857265.1 GI:2945567
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 49)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.

TITLE and Wright, D., Weiss, R.
 Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 JOURNAL Unpublished (2000)
 COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0266 row: 0 column: 11
 Seq primer: CGTCTAAGACGACGCCACAT
 Class: plasmid ends
 High quality sequence stop: 60.

FEATURES

Source

1. 60
 Location/Qualifiers
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUC1M0266011"
 /clone_lib="Mouse 10kb plasmid UUC1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g114732114|g1b|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT

6 a 16 c 23 g 15 t

ORIGIN

Query Match 50.4%; Score 12.6; DB 13; Length 60;
 Best Local Similarity 78.9%; Pred. No. 1.1e+05;
 Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 4 gctcgcccatcaatacata 22
 || ||||| |||||
 Db 23 gccggcccccacatcacata 5

RESULT 15

AZ331596 36 bp DNA GSS 29-SEP-2000
 LOCUS 1M0059E08R Mouse 10kb plasmid UUC1M library Mus musculus genomic
 DEFINITION clone UUC1M0059E08 R, DNA sequence.
 ACCESSION AZ331596
 VERSION AZ331596.1 GI:10394441
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 36)
 AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,
 Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly

TITLE 'M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
 and Wright, D., Weiss, R.
 Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 JOURNAL Unpublished (2000)
 COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0053 row: E column: 08
 Seq primer: CACACGAGAAACGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 36.

FEATURES

Source

1. 36
 Location/Qualifiers
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUC1M0059E08"
 /clone_lib="Mouse 10kb plasmid UUC1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g114732114|g1b|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT

13 a 9 c 0 g 14 t

ORIGIN

Query Match 49.6%; Score 12.4; DB 13; Length 36;
 Best Local Similarity 72.7%; Pred. No. 1.3e+05;
 Matches 16; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 1 acagctcgcccatcaatacata 22
 ||| ||||| ||||| |||||
 Db 5 ACATCTCCCCCTTTTAAATATA 26

Search completed: March 9, 2002, 00:09:21
 Job time: 11037 sec

THIS PAGE BLANK (USPTO)
